What you submitted
Feb 22

Title

Global PII leak on IG - Minors' info leaked and 7% of EU users' phone #s revealed 26% of their email addresses

Vuln Type

Identification / Deanonymization

Product Area

Instagram

Description/Impact

Description
===
PII is literally embedded in the source code of every user's profile page. This is especially significant for those accounts whose profile type is "PERSON" and who have NOT set their account to private mode. Approximately 60% of all IG accounts fall within the universe of users at-risk for leak of PII.

This PII is readily viewable and can be accessed by anyone who is logged into IG.

Primarily, the PII that is leaked is a person's private email address and less frequently, the user's phone #, city/state/province and portions of their address (less often street name and number, but city/state/region/country information has been leaked for some users)

I've reviewed over 59,000 IG profiles with approximately 1/2 from the US and the other 1/2 are user living in the EU.

Based on my analysis:
** The phone #s of over 7% of EU based users has been leaked
** The email addresses of over 26% of EU based users has been leaked

Here are several specific examples:

MINORS IN THE EU WHOSE PII HAS BEEN LEAKED

Example 1:
[redacted]
User is a minor - he has self-identified as 14 years old
PII leaked:
phone number
Email address
City, region and country
Per schema.org code, this user's account is a 'personal' profile
A google search for his email address finds zero results implying that he has not posted it where others could readily find it.

Example 2:

[redacted]
User is a minor – she indicates on her profile that she's "14y" and she may have posted her birthdate coded in roman numerals as "IX.VIII.MMIV" (August 4, 2004)
PII leaked:
Email address
Her email is found on source code page and the format of her email address allows one to personally identify her.
A quick google search for the 2 names in her email address confirm that her true name can be derived from her email

[redacted] to confirm this

Example 3:
[redacted]
User is a minor – he indicates on his profile that he is "14y/o"
PII leaked:
Email address

Similar to the first example, a google search for this user's email address finds zero results

Analysis of PII leak among 23,994 EU based IG users

I reviewed the profiles of 23,994 IG users who are highly likely to be living in the EU.

These users were located by two primary methods:
Method 1:
I Identified 10 IG users who live in the EU and compiled a list of their followers (see note 1). This method yielded a total of 12,119 user profiles, of which 10,510 were personal pages (not businesses or organizations)

Here are the number of IG users with a Personal profile page (not a business or organization) broken out by the value of their privacy flag:

[redacted]

The relatively high percentage of users who have flagged their accounts as private is likely due to the fact that many of these followers were teenagers

Method 2:
To get a more geographically representative sample, I compiled a list of the user names for those posts which are featured on IG's 'Explore Locations' pages.

A total of 1,764 different location pages were reviewed.
Locations were limited to those within Germany and Austria.
These pages were found by getting the urls for the country specific pages listed
within https://instagram.com/explore/locations/

Each page had 33 posts on it for a total of 57,984 posts

I would like to submit a file that contains the url of the location page and the post ID #s for the posts which were on the page at the time I visited.

Once I had the list of posts, I then reviewed the content of 29,497 separate posts from which I compiled the user names of the original poster as well as the names of the users who had most recently commented or liked the post.

I then reviewed the profile pages for those individuals and ended up with 12,526 profiles that were not flagged as private and whose profile is of type "PERSON"

Impact
===
As noted above, the fact that PII is 'hiding in plain sight' means that virtually anyone could scrape this information for tens of millions of users.

The phone #s of over 7% of EU based users has been leaked
The email addresses of over 26% of EU based users has been leaked

I have also identified minors living in the EU whose PII has been leaked and because this is a systemic issue there could be thousands more instances of minors' PII being leaked. I would like to provide a list of 22 users who appear to be minors. I have identified thes following 26 IG users as likely minors living in the EU and of those, 3 have a leaked phone # and 8 have leaked email address:

[redacted]

Because this personal information is hiding in plain sight, FB and IG could suffer severe negative PR were this to be disclosed

As a person who has been datamining on the web (wearing a white hat) I was especially surprised to see that this leak existed and that it was so blatantly clear to see if anyone looked for it.

[redacted]

I would like to send over documentation that further demonstrates the scope of this issue.

Repro Steps

Setup
===
Go to any user's profile page and view the source code

That's all there is to it !!!

Seriously!

It appears that the LD+JSON link is the culprit and the fact that even though the profiles are for private people, the PII fields all start with "business_" (eg business_phone, etc)

Please let me know how I can provide you with my documentation and research.

Thank you,

David

[redacted]

Feb 22
Hi,

Thank you for reporting a security issue! Your report number is 3022413804650803. Please give us reasonable time to

investigate and mitigate the issue before sharing information with others, and note that we reserve the right to publish your report. (More details: https://www.facebook.com/whitehat/) Note that if you're writing to us in a language other than English, we'll only able to respond in English at this time. We're sorry for any inconvenience this causes.

If you're trying to report another issue, please review the information below to get help.

- If your account or a friend's account is sending out suspicious links: https://www.facebook.com/help/hacked
- To report abuse: https://www.facebook.com/help/reportlinks
- To report bugs that are not security issues: https://www.facebook.com/help/www/326603310765065
- For any other questions or concerns, please visit our Help Center: https://www.facebook.com/help

Thanks,
Facebook Security

Your reply
Feb 24
One more item that I wanted to share is the attached table that documents the percent of all EU IG users broken down by the PII leaked, the profile type and the user's privacy setting. 4.7% of users in the EU have likely had their phone # leaked and 16.9% have had their email address leaked. These percentages are different from the ones in my initial subject line because these new percentages include business accounts and organizations - so these percentages are truly the percent all IG users in the EU.
Attachments
Percent of EU users with PII data leak.png

Your reply
Today
Two items to note:
1) Although my report focuses on the impact of this data leak in the European Union, I have also analyzed the profiles of nearly 30,000 US based users and the extent of the data leak for US users is on a comparable scale to that of users in the European Union. I've attached a table that documents these numbers.

2) I have not yet heard back from someone and given the global scale of this data leak and the fact that millions of users have had their personally identifiable information revealed without their consent, I am adding the following keywords to see if that will speed up the process of your review of my report:

keywords:
GDPR,
EU,
European Union member countries
data breach,
global,
millions,
privacy,
minors
PII
Personally identifiable information
Public relations
PR
legal

I hope to hear from someone soon.

Thank you
Attachments
Percent of USA users with PII data leak.png

Note: The date shown in the reply above is "Today". This reply was sent on February 28 at approximately 8:45AM

Mar 7
Hi David,

Thank you for your report! From looking at the information you provided, we've confirmed that the accounts you've identified here correspond to people who have turned their profiles into Business Profiles. That is why the fields you mentioned here were prefixed with "business_". https://help.instagram.com/138925576505882contains some background on what Business Profiles are and https://help.instagram.com/502981923235522 describes how someone can enable / disable Business Profile features on their account. The process of enabling Business Profile features requires a person to explicitly opt in.

As that first link notes, Business Profiles can provide a public "Contact" button to their profile using information supplied by the profile owner. This contact information is what you saw in the HTML of the page. People always have the ability to change their contact information via account settings. Beyond that, during the setup process for Business Profiles we display this information, remind people that it will be accessible to others, and allow them to update or remove the information. After discussing this functionality with the Instagram team we did take steps to remove the contact information from the HTML of the page, since it was not necessary to include in its current form. However this information is still accessible to Instagram users via the Contact button.

Given our assessment, this report would not qualify under our program. Please let me know if you have any additional questions here.

Thanks,

Neal
Security

Neal,

Thanks for your reply.

Unfortunately there are several significant errors in your assessment.

1) *Business versus Personal Profiles*
As I noted in my report, I determined the type of profile each user had by examining the string of code found in the source code of each user. Here's a real example of that source code:

I've highlighted the @type field above and I labeled all profiles whose source code for the field "@type" had a value of *"Person" *as being **not** a business but instead as a person. In my survey of EU IG users, there were *over
18,000 profiles of private individuals* (value of "Person" in @type field)

I specifically identified other non-personal profiles (such as a business profile). For my report, I created a group identified as *"All Others"* * *only** if the value for the field "@type" was among the following values:

AutomotiveBusiness
FoodEstablishment
GovernmentOrganization
GroceryStore
HealthAndBeautyBusiness
HomeGoodsStore
LocalBusiness
Store

There were only 1,286 profiles which I grouped as having a profile type among those listed above.

There were also 1,395 profiles with a value of "Organization".

There was no data in the source code to identify the user type for 1,813 users.

2) Even if your explanation was correct that over 18,000 people changed their profile type to that of a business please explain the following:
a) the email was NOT visible on the profile pages of those 18,000 people nor was there a "contact" button on any of the pages that I viewed. I can send the source code of many users' profiles so that you can verify this for yourself.
b) the source code clearly identifies their profile as being for a "Person" so how could someone have changed their profile to a business profile yet have nothing in the source code of their page which indicates that they have a business profile?

I look forward to your reply.

Thank you,

David

David J. Stier
Marketing Executive | Data Scientist
[redacted]

linkedin.com/in/davidjstier
realworlddatascience.com

My Data Science Presentation ranked in Top 10 (of 800) in Bay Area last year:
http://bit.ly/Top10_DataSci_talks

On Thu, Mar 7, 2019 at 5:06 PM Facebook <
case++aazqhnizwcnq7e@support.facebook.com> wrote:

Neal,

I've uploaded all the data that I used for generating my reports onto google drive which you can access here:

https://drive.google.com/file DELETED
If your team audits the profiles of 50 users who I've identified as having type "Person", I am sure you will see that these profiles were never set up as business profiles.

I have also placed a copy of the webpages for selected profile pages that I saved on March 1 which clearly show the presence of PII in the source code. You can access those profiles here:
https://drive.google.com/drive/DELETED

Thank you,

David

David J. Stier
Marketing Executive | Data Scientist
[redacted]

linkedin.com/in/davidjstier
realworlddatascience.com

My Data Science Presentation ranked in Top 10 (of 800) in Bay Area last year:
http://bit.ly/Top10_DataSci_talks

On Thu, Mar 7, 2019 at 5:06 PM Facebook <
case++aazqhnizwcnq7e@support.facebook.com> wrote:

Hi David,

Thanks for following up. Let me see if I can address your concerns here.

> 1) *Business versus Personal Profiles*

>I've highlighted the @type field above and I labeled all profiles whose
> source code for the field "@type" had a value of *"Person" *as being **not**
> a business but instead as a person.

I double-checked the code here to confirm and I can verify that the @type that we displayed there would be "Person" for personal profiles, and also for some sets of business profiles (ie: those representing celebrities). The contact information would only be displayed for business profiles.

> a) the email was NOT visible on the profile pages of those 18,000 people
> nor was there a "contact" button on any of the pages that I viewed. I can
> send the source code of many users' profiles so that you can verify this
> for yourself.

As I mentioned to you earlier, Business Profiles can provide a public "Contact" button to their profile using information supplied by the profile owner. This contact information is what you saw in the HTML of the page. People always have the ability to change their contact information via account settings. Beyond that, during the setup process for Business Profiles we display this information, remind people that it will be accessible to others, and allow them to update or remove the information.

After discussing this functionality with the Instagram team we did take steps to remove the contact information from the HTML of the page, since it was not necessary to include in its current form. However this information is still accessible to Instagram users via the Contact button in our mobile application.

> b) the source code clearly identifies their profile as being for a
> "Person" so how could someone have changed their profile to a business
> profile yet have nothing in the source code of their page which indicates
> that they have a business profile?

As I mentioned, certain types of "business" profiles representing people would be identified as a "person" in the data you saw.

I took a look at the data you provided here and all of that seems consistent with the explanation I've provided you. Please let me know if you have any additional questions here.

Thanks,

Neal
Security
Neal,

The PII was in the HTML code of more than 18,000 profiles that were of type "Person".

This was NOT a case of some celebrity pages being considered a personal page.

If you just look at 10 or so of the profiles I've provided, you can see that these indeed are personal pages.

I think that you are confusing the fact that the LD+JSON had the variable "business_email" when the profile for the page being shown was that of a person. How that "business" variable got into the HTML of more than 18,000 personal profiles is the real issue.

If you start with the fact that these 18,000+ profiles are indeed personal profile pages [which you agree with in item 2 above] then the business_email should not appear (in fact no email whatsoever should appear).

However, in your approach you've started with the fact that there's a variable that could only appear on a business profile page so therefore it must be a business profile page - this logic doesn't hold up.

Thank you,

David

> The PII was in the HTML code of more than 18,000 profiles that were of type "Person".
>
> This was NOT a case of some celebrity pages being considered a personal page.
>
> If you just look at 10 or so of the profiles I've provided, you can see that these indeed are personal pages.

Again, any profile can opt in to being a business profile. When they do so, they select the type of business their account represents. That includes options which represent people (ie: Personal Chef, for example). Those options would cause us to render a "type" of Person in the data you saw, since the profile represents a person. That field does not indicate whether or not business profile features have been enabkl

> I think that you are confusing the fact that the LD+JSON had the variable "business_email" when the profile for the page being shown was that of a person. How that "business" variable got into the HTML of more than 18,000 personal profiles is the real issue.

There is no confusion on my end. I have explained to you why you observed the behavior that you did. The fact that you saw the business_email field indicates that these *were* accounts that had enabled the business profile feature because the field was only displayed for accounts which enabled these features.

Thanks,

Neal
Security

Neal,

Thank you for your reply.

Given your explanation, it sounds like Instagram would not be concerned if I were to call or text message all the individuals whose phone number was revealed to let them know that their personally identifiable information was revealed by Instagram and that Instagram does not consider this to be any problem because every one of these people had changed their profile type to that of a business profile.

I look forward to finding out if your users in the EU agree with your determination.

David

On Mon, Mar 11, 2019, 4:14 PM Facebook <
case++aazqhnizwcnq7e@support.facebook.com> wrote:

Hi David,
As I've explained to you here, the set of information that you observed corresponds to accounts which have enabled Instagram business profiles and have opted to provide contact information. I'd encourage you to look at these profiles in the Instagram app itself to see that they have contact links corresponding to this information. For example, the @zweihochfuenf profile you mentioned earlier has an "email" link which points to the business email they provided.
Although this issue does not qualify as a part of our bounty program we appreciate your report. We will follow up with you on any security bugs or with any further questions we may have.
Thanks,

Neal
Security

https://www.instagram.com[redacted]

Instagram

Search

[redacted]

For 18,022 EU Instagram Users analyzed which specific PII is leaked?

# of EU IG users Phone leak?	Email leak?	Srt Addr leak?	City leak?	Method used EU Locations	IG Followers	Grand Total
yes	**no**	yes	yes	4		4
		no	no	100	77	177
	yes	yes	yes	91	16	107
			no	1		1
		no	no	803	261	1,064
yes Total				**999**	**354**	**1,353**
no	**no**	yes	yes	17	5	22
		no	no	8,382	4,537	12,919
	yes	yes	yes	47	1	48
		no	no	3,081	599	3,680
no Total				**11,527**	**5,142**	**16,669**
Grand Total				**12,526**	**5,496**	**18,022**

NOTE - all users analyzed above have their 'is private' flag set to FALSE and a profile type value of "PERSON"

KEY FINDINGS:

7.5% of all users' phone #s have been leaked (1,353 of 18,022)

26.6% of all users' email addresses have been leaked (4793 of 18022)

71.7% of all users have had NO PII data leak (12,919 of 18,022)

Image above: *KKeyyfindings png*

[redacted]

4.7% of all IG accts in EU have leaked phone # (orange and light blue cells)
12.9% of all IG accts in EU have leaked email address (light yellow cells)

# of EU IG users Accts Privacy Se	Profile Type	Phone leak?	Email leak?	Srt Addr leak?	City leak?	Method used EU Locations	IG Followers	Grand Total
PUBLIC	Unknown	no	no	no	no	4.3%	1.9%	6.3%
	Unknown Total					4.3%	1.9%	6.3%
	All others					3.9%	0.5%	4.4%
	Organization					3.9%	0.9%	4.8%
	Person	yes	no	yes	yes	0.0%	0.0%	0.0%
				no	no	0.3%	0.3%	0.6%
			yes	yes	yes	0.3%	0.1%	0.4%
					no	0.0%	0.0%	0.0%
				no	no	2.8%	0.9%	3.7%
		no	no	yes	yes	0.1%	0.0%	0.1%
				no	no	29.0%	15.7%	44.7%
			yes	yes	yes	0.2%	0.0%	0.2%
				no	no	10.7%	2.1%	12.7%
	Person Total					43.3%	19.0%	62.3%
PUBLIC Total						55.5%	22.4%	77.9%
PRIVATE	Unknown	no	no	no	no	0.5%	2.1%	2.6%
	Unknown Total					0.5%	2.1%	2.6%
	Person	no	no	no	no	3.3%	16.2%	19.5%
	Person Total					3.3%	16.2%	19.5%
PRIVATE Total						3.8%	18.3%	22.1%
Grand Total						59.3%	40.7%	100.0%

Distribution of all EU IG users based on privacy setting, type of profile and PII leak impact
Based on analysis of 28,912 users © David J Stier

Image above: *Percent of EU users with PII data leak png*

EXTENT OF DATA LEAK ON US BASED IG USERS

Distribution of 28.8k US based IG user profiles analyzed

Email leaked for 6.81% of ALL US based IG users and 11.4% of all "Person" profiles

Phone leaked 3.0% of ALL US based IG users and 3.8% of all "Person" profiles

Profile Type	Phone leak?	Email leak?	Total
Person	No	No	69.81%
	No	Yes	**6.40%**
	No total	either	76.21%
	Yes	No	**0.41%**
	Yes	Yes	**2.63%**
	Yes total	either	3.04%
Person Total			79.25%
Organization	No	No	0.85%
	No	Yes	2.59%
	No total	either	3.44%
	Yes	No	0.25%
	Yes	Yes	1.81%
	Yes total	either	2.05%
Organization Total			5.50%
Unknown	No total	No	9.38%
	No total	No	9.38%
Total			9.38%
All others	No	No	0.90%
	No	Yes	2.10%
	No total	either	3.01%
	Yes	No	0.27%
	Yes	Yes	2.60%
	Yes total	either	2.87%
All others Total	All	All	5.87%
Grand Total			100.00%

Image above: *Percent of USA users with PII data leak png*